

-----X

ULTIMA SERVICES CORPORATION,

Plaintiff,

-against-

No.2:20-cv-00041
DCLC-CRW

U.S. DEPARTMENT OF AGRICULTURE,
U.S. SMALL BUSINESS ADMINISTRATION,
SECRETARY OF AGRICULTURE, and
ADMINISTRATOR OF THE SMALL BUSINESS
ADMINISTRATION,

Defendants.

-----X

Jonathan Guryan hereby states:

Case 2:20-cv-00041-DCLC-CRW Document 58-6 Filed 06/21/22 Page 1 of 4
PageID #: 407

Economics Department and the Kellogg School of Management at Northwestern University.

2. In connection with this case I reviewed the report of Defendants' expert Daniel Chow. Mr. Chow's report stated in pertinent part:

From the original raw datasets, I compiled a list of relevant variables from both SAM registrants' data (7,466,447 observations and 42 variables) and FPDS awards (5,104,224 observations and 55 variables).

* * *

Merging and removal of redundant and extraneous observations by DUNS number resulted in a *combined file of 5,659,740 registration observations and 64 variables.* (emphasis added)

3. I advised Plaintiff to request Mr. Chow's data files so that I could reproduce and evaluate Mr. Chow's analysis.

4. I was given two of the files Mr. Chow used. One is named "SAM_1904_2009_SB_D.dta" (7,466,447 obs and 42 variables), and the other is "Awards_Combined_FY19FY20_SBNonSB.dta" (5,659,740 obs and 64 variables). The "Awards_Combined" file contained the exact same number of observations and variables as the last sentence quoted *supra* in the excerpt from the Chow Report. Given the reference to it at the end of the quoted excerpt, in a paragraph largely about SAM data, I assumed that the latter file was a combination of the FPDS and SAM data, and nothing in the Chow Report suggested otherwise.

5. At my request, Plaintiff asked Mr. Chow during his deposition whether the 64 variable, 5.6 million observation file he referred to in his report is the combined FPDS-SAM file. Mr. Chow confirmed that it is. I therefore based much of my report on this assumption, and identified the 64 variable 5.6 million observation “Awards Combined” file as Mr. Chow’s “analysis” file, as Mr. Chow had testified, combining the data from the SAM and FPDS files.

6. Shortly after March 17, 2022, I reviewed Mr. Chow’s deposition transcript along with the two data files Defendants produced. I determined it would be impossible to replicate Mr. Chow’s analysis with the files we had been given, and the information in Mr. Chow’s deposition transcript. I asked for and eventually received the computer code Mr. Chow used.

7. The computer code produced by Defendants refers to a file called “FPDS_SAM_PPIRS_2019_2020_matching_Test4.” The code indicates that this “matching file” is used to create the “finalize” file that Mr. Chow utilized for his analysis to reach his conclusions. However, the derivation of the “matching file” is not clear.

8. In reviewing Mr. Chow’s computer code, I determined that, in all other instances, the end result of one program is a file that is used at the outset of the next program, sequentially. But the program prior to the use of the “matching” file program does not end with the “matching” file. Further, the “matching” file has a key variable – “winner” – but there is no code

showing how Mr. Chow created that variable or how it was determined whether a given company was a “winner.”

9. On May 17, 2022, Defendants produced a “matching file,” named “FPDS_SAM_PPIRS_2019_2020_matching_Test4.” (having 1,259,129 obs and 50 variables).

10. Mr. Chow used STATA software in performing his analysis. In STATA, the term “collapse” is a term of art that can refer to a wide variety of different actions. Without knowing exactly what Mr. Chow did to “collapse” the data, it is impossible to replicate his work. There is no reference in Mr. Chow’s report to his “collapsing” data when combining the FPDS and SAM data. Indeed, there is no reference in Mr. Chow’s report to a file that has 1,259,129 observations and 50 variables.

I state under penalty of perjury that the foregoing is true and correct.

Executed on June 21, 2022



Dr. Jonathan Guryan